

ИСПОЛЬЗОВАНИЕ СЕГМЕНТАЦИИ ДЛЯ АВТОМАТИЗАЦИИ ДЕШИФРИРОВАНИЯ МНОГОСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ

© 2009 г. Р. М. Алеев*, доктор техн. наук; В. Б. Фофанов**, канд. техн. наук

* ПО “УОМЗ”, филиал “УОМЗ–Институт прикладной оптики”, г. Казань

E-mail: optool@tbit.ru;

** Казанский государственный университет, г. Казань

E-mail: Viatcheslav.Fofanov@ksu.ru

Исследуются возможности сегментации зон интереса по многоспектральным изображениям применительно к двум математическим моделям сцены. Показано, что применение необоснованных предположений при построении модели может привести к получению ложных сведений о сцене. Приведены результаты компьютерных экспериментов по сегментации зон интереса с использованием трехспектральных изображений.

Ключевые слова: автоматизация дешифрирования, модели сцены, зоны интереса, многоспектральные изображения, сегментация сцены.

Коды OCIS: 100.0100.

Поступила в редакцию 18.03.2009.

Введение

Обязательным этапом автоматизации дешифрирования является создание его математической модели. К сожалению, несмотря на несомненные успехи, достигнутые в распознавании печатных символов и идентификации людей по папиллярным узорам, достаточно общей теории дешифрирования изображений, необходимой для его автоматизации, пока построить не удалось. Настоящая работа касается частного, но распространенного случая, когда дешифрирование проводится с целью выявления (обнаружения) на сцене заданных объектов.

Понятно, что выявляемые объекты должны обладать характерными свойствами (признаками), позволяющими отличать их от других объектов сцены. Также очевидно, что эти признаки, хотя бы частично, должны присутствовать на изображениях. Однако измеряемые в процессе съемки энергетические яркости пикселов очень изменчивы. Они зависят от времени суток, сезона, метеопараметров и других трудно контролируемых условий съемки. Гораздо устойчивее ведут себя геометрические признаки (форма, площадь, габаритные размеры). Именно с этим свойством связано их широкое применение в дешифрировании изображений. К сожалению, геометрические признаки не регистрируются в

ходе съемки, а вычисляются по проекциям объектов на этапе дешифрирования. Поэтому построение проекций заданных объектов является обязательным этапом дешифрирования.

Важная особенность рассматриваемого подхода к выявлению на сцене заданных объектов заключается в том, что эту задачу предлагается заменить тремя более простыми задачами. Вначале изображения сцены используются для поиска на ней участков, содержащих заданный объект и его некоторое окружение. Такие участки названы зонами интереса. Затем проводится сегментация выявленных зон интереса, состоящих из пикселов объекта и его окружения (фона). Целесообразность введения понятия зоны интереса оправдывается двумя соображениями. Во-первых, сегментировать приходится только выявленные зоны интереса. Если частота обнаружения ложной зоны невелика, то их общая площадь оказывается во много раз меньше площади всей сцены. Во-вторых, естественно полагать, что качество сегментации зоны интереса, содержащей пиксели только двух видов, окажется значительно выше качества сегментации всей сцены. Координаты пикселов, образующих объект, называются далее его проекцией. На завершающем этапе проекции используются для вычисления геометрических признаков и принятия окончательного решения о наличии объекта в зоне интереса.

Очевидно, что состав признаков, которые можно вычислять по изображениям, определяется выбранной математической моделью сцены. Стремление к получению более полной информации о сцене может привести к использованию моделей, построенных на непроверенных предположениях. При решении прикладных задач такие предположения нередко приводят к неверным выводам. В настоящей работе приводятся результаты теоретических и экспериментальных исследований влияния на результат поиска заданных объектов разных моделей сцены. Из-за ограниченного объема статьи рассматривается только задача сегментации зон интереса. В качестве исходной информации о сцене используется набор пространственно совмещенных и одновременно сформированных изображений.

Модели сцены

Будем рассматривать сцену как совокупность неделимых элементов, называемых далее пикселями. Предположим вначале, что в каждом пикселе измеряется только один признак. В этом случае пиксель характеризуется целочисленными координатами $z = (z_1, z_2)$, заданными на двумерной целочисленной решетке $Z^2 = \{z = (z_1, z_2) : z_1 \in Z, z_2 \in Z\}$, и скалярной случайной величиной ξ_z со значениями из конечного множества Y , состоящего из $|Y| > 1$ элементов. Предполагается, что $Y = \{0, 1, \dots, |Y| - 1\}$ и что случайные величины определены на одном и том же вероятностном пространстве (Ω, A, P) . Очевидно, что в качестве Ω можно рассматривать счетное множество Y^{Z^2} всех отображений вида $Z^2 \rightarrow Y$, а в качестве σ -алгебры – A – счетное семейство всех подмножеств множества Y^{Z^2} . Семейство вида $(\xi_z)_{z \in Z^2}$ будет называться далее скалярной сценой. Пусть $\omega \in \Omega$, $z \in Z^2$ и $x_z = \xi_z(\omega)$. Отображение $x : Z^2 \rightarrow Y$, определяемое равенством $x(z) = x_z$, $z \in Z^2$ и обозначаемое $x = (x_z)_{z \in Z^2}$, назовем скалярным изображением сцены.

При решении прикладных задач интерес представляют конечные подмножества пикселов, которые будут называться объектами. Формально каждый объект определяется подмножеством A точек из Z^2 , содержащим координаты его пикселов, и семейством $\xi_A = (\xi_a)_{a \in A}$ из $|A|$ скалярных случайных величин. Далее A будет называться проекцией объекта. Если A и B – проекции разных объектов, то предполагается, что они не пересекаются. Изображением объекта ξ_A будет называться сужение $x_A = (x_a)_{a \in A}$ изображения x всей сцены на A .

Пусть $Y^A = \{x_A = (x_a)_{a \in A} : x_a \in Y, a \in A\}$ – множество различных изображений объекта ξ_A , тогда его свойства определяются распределением вероятностей $P_{Y^A} = (p_{Y^A}(x_A))_{x_A \in Y^A}$ на Y^A . Если на Z^2 задано разбиение, состоящее из конечных попарно непересекающихся подмножеств, называемых проекциями объектов сцены, и если каждой проекции A поставлено в соответствие распределение вероятностей P_{Y^A} на Y^A , то существует [1] вероятностное пространство (Ω, A, P) и скалярная сцена $(\xi_z)_{z \in Z^2}$ на Z^2 такая, что

$$P\{\omega \in \Omega : \xi_A(\omega) = x_A\} = p_{Y^A}(x_A)$$

для любой проекции A и для любого $x_A \in Y^A$. Кроме того, если A и B – проекции разных элементов сцены, то для любых $a \in A$ и $b \in B$ случайные величины ξ_a и ξ_b независимы.

В общем случае число одновременно измеряемых признаков $v \geq 1$. Поэтому в качестве пикселя с координатами $z \in Z^2$ естественно рассматривать v -мерную случайную величину $\xi_z = (\xi_z^j)_{1 \leq j \leq v}$, определенную на (Ω, A, P) и принимающую значения в $Y = Y^v$, а семейство $(\xi_z)_{z \in Z^2}$ – называть векторной сценой. Пусть $\omega \in \Omega$, тогда $x_z^j = \xi_z^j(\omega)$, $z \in Z^2$, будет называться j -м скалярным изображением пикселя ξ_z , а $x^j = (x_z^j)_{z \in Z^2}$ – j -м скалярным изображением векторной сцены, $1 \leq j \leq v$. Изображением векторной сцены (или векторным изображением) назовем семейство $x = (x^j)_{1 \leq j \leq v}$ ее скалярных изображений. Если $x_z = (x_z^j)_{1 \leq j \leq v}$ – векторное изображение пикселя, то изображение векторной сцены принимает вид $x = (x_z)_{z \in Z^2}$.

Объектом с проекцией A в общем случае будет совокупность векторных случайных величин $\xi_A = (\xi_a)_{a \in A}$, а его изображением – совокупность векторов $x_A = (x_a)_{a \in A}$. Очевидно, что изображения объекта образуют множество

$$Y^A = \{x_A = (x_a)_{a \in A} : x_a = (x_a^j)_{1 \leq j \leq v}, a \in A\}.$$

Распределением объекта ξ_A назовем распределение $P_{Y^A} = (p_{Y^A}(x_A))_{x_A \in Y^A}$ на множестве Y^A всех векторных изображений, определяемое равенством $p_{Y^A}(x_A) = P\{\omega \in \Omega : \xi_A(\omega) = x_A\}$. Пусть на Z^2 задано разбиение на проекции и пусть каждой проекции A поставлено в соответствие распределение вероятностей P_{Y^A} на множестве Y^A . Тогда существует вероятностное пространство (Ω, A, P) и векторная сцена $(\xi_z)_{z \in Z^2}$ такая, что

$$P\{\omega \in \Omega : \xi_A(\omega) = x_A\} = p_{Y^A}(x_A)$$

для любой проекции A и любого $x_A \in Y^A$. Кроме того, если A и B – проекции разных элементов сцены, то для любых $a \in A$ и $b \in B$ векторные случайные величины ξ_a и ξ_b независимы.

В ходе дешифрирования признаки объектов вычисляются по изображениям сцены. Пусть d – евклидово расстояние на Z^2 , а $B(\mathbf{z}, r) = \{\mathbf{t} \in Z^2 : d(\mathbf{z}, \mathbf{t}) \leq r\}$ – круг с центром \mathbf{z} и радиусом r . Довольно часто в качестве признака объекта скалярной сцены используется среднее арифметическое значение $\bar{x}_{\mathbf{z}}$, определяемое равенством

$$\bar{x}_{\mathbf{z}} = \frac{1}{|B(\mathbf{z}, r)|} \sum_{\mathbf{t} \in B(\mathbf{z}, r)} x_{\mathbf{t}}.$$

В общем случае $\bar{x}_{\mathbf{z}}$ не имеет полезной содержательной интерпретации. Однако в частном случае, если объект ξ_A со средним значением m_A является фрагментом изотропного случайного поля, для которого выполняются условия эргодической теоремы Слуцкого, то $\bar{x}_{\mathbf{z}}$ является состоятельной оценкой неизвестного среднего значения m_A [2]. Пусть R – корреляционная функция изотропного случайного поля со средним значением m_A . Хорошо известно [2], что для сходимости $\bar{x}_{\mathbf{z}}$ к m_A по вероятности достаточно, чтобы $R(t) \rightarrow 0$ при $t \rightarrow +\infty$. Это условие часто используется в приложениях.

В связи с изложенным будем называть далее скалярную сцену локально изотропной, если ее объекты являются фрагментами скалярных изотропных случайных полей, позволяющих оценивать по изображению объекта его среднее значение. Очевидно, что векторная сцена $(\xi_{\mathbf{z}})_{\mathbf{z} \in Z^2}$ является набором из v скалярных сцен $(\xi_{\mathbf{z}}^j)_{\mathbf{z} \in Z^2}$, $1 \leq j \leq v$. Она будет называться локально изотропной, если все эти скалярные сцены являются локально изотропными. Из этого определения следует, что $E\xi_{\mathbf{a}}^j = m_A^j$, $\mathbf{a} \in A$ и $1 \leq j \leq v$. Вектор $\mathbf{m}_A = (m_A^j)_{1 \leq j \leq v}$ будет называться вектором средних значений объекта ξ_A . Для соседних объектов ξ_A и ξ_B всегда предполагается, что $d(\mathbf{m}_A, \mathbf{m}_B) > 0$.

Довольно часто, кроме среднего арифметического значения, по изображениям вычисляются и другие признаки. В этом случае, помимо локальной изотропности к случайному полю, необходимо предъявлять дополнительные требования. Если предположить, что образующие объект ξ_A случайные величины в совокупности независимы и имеют одно и то же распределение $P_A = (p_A(y))_{y \in Y}$, то изображение x_A превращается в случайную выборку. В этом случае его можно использовать для оценки неизвестных вероятностей $p_A(y)$, $y \in Y$, и числовых характеристик распределения P_A , включая среднее значение m_A . В противном случае, у относительной частоты появления значения $y \in Y$ на изображении x_A отсутствует содержательная интерпретация.

Далее сцены, состоящие из объектов, образованных в совокупности независимыми случайными величинами с одним и тем же распределением, будут называться бернуlliевскими скалярными сценами. В случае векторных сцен предположение о независимости и равенстве распределений относится к векторным случайным величинам ξ_a , $a \in A$, образующим объект ξ_A . Очевидно, что каждая бернуlliевская сцена является локально изотропной.

Визуальный анализ скалярных изображений, полученных в различных спектральных зонах, свидетельствует о том, что средняя яркость заданного объекта часто оказывается выше или ниже средней яркости окружающего его фона хотя бы на одном из изображений. Объекты с таким свойством иногда называются пятнами. Очевидно, что пятна составляют довольно широкий класс объектов и поэтому заслуживают определенного внимания. Далее предполагается, что все заданные объекты являются пятнами.

Исследование пятен начнем с их формального определения. Пусть $Q(\mathbf{z}, l)$ – квадрат на Z^2 со стороной l и левым верхним пиксели $\mathbf{z} \in Z^2$, задаваемый равенством вида

$$Q(\mathbf{z}, l) = \{\mathbf{t} = (t_1, t_2) \in Z^2 : z_j \leq t_j \leq z_j + l, j = 1, 2\}.$$

Точки \mathbf{z} и \mathbf{t} из Z^2 будут называться соседями, если евклидово расстояние между ними $d(\mathbf{z}, \mathbf{t})$ равно единице. Очевидно, что каждый пиксель имеет четырех соседей. Подмножество $Fr(Q)$ точек из $Q(\mathbf{z}, l)$ назовем границей квадрата, если каждая точка $\mathbf{z} \in Fr(Q)$ имеет хотя бы одного соседа из $Z^2 \setminus Q$. Очевидно, что $|Fr(Q(\mathbf{z}, l))| = 4l$.

Рассмотрим частный случай, когда каждый пиксель $\mathbf{z} \in Z^2$ описывается скалярной случайной величиной $\xi_{\mathbf{z}}$, а исходной информацией о сцене служит единственное изображение $x = (x_{\mathbf{z}})_{\mathbf{z} \in Z^2}$. Пусть A – связное конечное подмножество из Z^2 . Объект ξ_A будет называться светлым (соответственно темным) пятном с диаметром $d(A)$, если, во-первых, существует квадрат Q на Z^2 такой, что $A \subset (Q \setminus Fr(Q))$; во-вторых, $E\xi_{\mathbf{z}} = m_{Q \setminus A}$, $\mathbf{z} \in Q \setminus A$, и, в-третьих, $m_A > m_{Q \setminus A}$ (соответственно $m_A < m_{Q \setminus A}$). Пиксели $\xi_{Q \setminus A} = (\xi_{\mathbf{z}})_{\mathbf{z} \in Q \setminus A}$ будут называться окрестностью пятна (или фоном).

В общем случае, когда $v \geq 1$, первое условие остается без изменения, второе условие принимает вид $E\xi_{\mathbf{z}} = m_{Q \setminus A}$, $\mathbf{z} \in Q \setminus A$, а третье – $d(\mathbf{m}_A, \mathbf{m}_{Q \setminus A}) > 0$, означает, что хотя бы для одной скалярной сцены объект является пятном.

При работе с векторными случайными величинами используется введенное ранее понятие

случайного расстояния [3]. Приведем его определение, которое понадобится далее.

Расстояние d на R^v будет называться борелевским, если борелевским является отображение $d: R^v \times R^v$. Из свойств борелевских функций следует, что расстояние Евклида и многие другие известные расстояния являются борелевскими. Пусть d – борелевское расстояние на R^v , а $\xi = (\xi_j)_{1 \leq j \leq v}$ и $\eta = (\eta_j)_{1 \leq j \leq v}$ – векторные случайные величины, определенные на некотором вероятностном пространстве (Ω, \mathcal{A}, P) . Известно, что отображение $d(\xi, \eta): \Omega \rightarrow R$, определяемое равенством $d(\xi, \eta)(\omega) = d(\xi(\omega), \eta(\omega))$, также будет случайной величиной на (Ω, \mathcal{A}, P) . Из определения расстояния следует, что при каждом $\omega \in \Omega$ будут выполняться аксиомы расстояния. Это позволяет назвать отображение d случайным расстоянием на множестве v -мерных случайных величин, а случайную величину $d(\xi, \eta)$ – случайным расстоянием между ξ и η .

Сегментация

Пусть A – связное подмножество на Z^2 с диаметром $d(A)$, которое является проекцией объекта $\xi_A = (\xi_a)_{a \in A}$ векторной сцены с вектором средних значений $\mathbf{m}_A = (m_A^j)_{1 \leq j \leq v}$. Для любого $l \geq d(A) + 2$ существует квадрат Q на Z^2 со стороной l такой, что $A \subset Q \setminus Fr(Q)$. Будем называть семейство $\xi_Q = (\xi_z)_{z \in Q}$ зоной интереса для объекта ξ_A , если пересечение проекции Q с проекцией любого другого заданного объекта является пустым множеством.

Очевидно, что каждый объект с признаком пятна имеет зону интереса. Некоторые подходы к формализации задачи поиска зон интереса для пятен и результаты ее решения обсуждались ранее [4]. Поэтому в настоящей работе речь пойдет только о сегментации зон интереса на две части с именами “объект” и “фон”. Вначале рассмотрим решение этой задачи для локально изотропной векторной сцены.

Пусть $\xi_A = (\xi_a)_{a \in A}$ – заданный объект с площадью $|A|$, $\xi_Q = (\xi_z)_{z \in Q}$ – его зона интереса, а $\mathbf{m}_A = (m_A^j)_{1 \leq j \leq v}$ и $\mathbf{m}_{Q \setminus A} = (m_{Q \setminus A}^j)_{1 \leq j \leq v}$ – векторы средних значений самого объекта и его фона. Из определения зоны интереса следует, что для любых $a \in A$ и $z \in Q \setminus A$ имеет место неравенство

$$\begin{aligned} d(E\xi_a, \mathbf{m}_{Q \setminus A}) &= d(\mathbf{m}_A, \mathbf{m}_{Q \setminus A}) > 0 = \\ &= d(E\xi_z, \mathbf{m}_{Q \setminus A}). \end{aligned} \quad (1)$$

Покажем, что его можно применить для сегментации зоны интереса. Пусть

$$\begin{aligned} \bar{\mathbf{x}}_z &= \frac{1}{|B((\mathbf{z}, r))|} \sum_{t \in B(\mathbf{z}, r)} \mathbf{x}_t, \mathbf{z} \in Q \setminus Fr(Q) \\ \text{и } \bar{\mathbf{x}}_{Fr} &= \frac{1}{|Fr(Q)|} \sum_{t \in Fr(Q)} \mathbf{x}_t \end{aligned}$$

– оценки неизвестного вектора $E\xi_z$ и неизвестного вектора $\mathbf{m}_{Q \setminus A}$ соответственно, а $d(\bar{\mathbf{x}}_z, \bar{\mathbf{x}}_{Fr})$ – расстояние между ними. Так как d непрерывно, то $d(\bar{\mathbf{x}}_z, \bar{\mathbf{x}}_{Fr})$ сходится по вероятности к $d(E\xi_z, \mathbf{m}_{Q \setminus A})$ при $r \rightarrow +\infty$ [5]. Если упорядочить $d(\bar{\mathbf{x}}_z, \bar{\mathbf{x}}_{Fr})$, $\mathbf{z} \in Q \setminus Fr(Q)$ по возрастанию, то при достаточно большом радиусе сглаживания r на $|A|$ последних местах будут находиться расстояния $d(\bar{\mathbf{x}}_a, \bar{\mathbf{x}}_{Fr})$, $a \in A$, соответствующие пикселям объекта.

Отметим, что в изложенном методе квантилей используются только средние арифметические значения $\bar{\mathbf{x}}_z$, которые являются оценками неизвестных $E\xi_z$. Поэтому метод применим для локально изотропных сцен.

Рассмотрим далее некоторые из возможных решений задачи сегментации в рамках бернуллиевской векторной сцены. Начнем со случая, когда $v = 1$. Если рассматривать яркость пикселя в качестве признака, то сегментацию зоны интереса можно рассматривать как задачу классификации ее пикселов на два класса – с номером 1 (объект) и номером 2 (фон). Известно, что для построения байесовского решающего правила $h^*: Y \rightarrow \{1, 2\}$ с минимальной вероятностью ошибки классификации $e(h^*)$ требуется знать априорные вероятности $P(A)$, $P(Q \setminus A)$ и распределения признаков $P_A = (p_A(y))_{y \in Y}$, $P_{Q \setminus A} = (p_{Q \setminus A}(y))_{y \in Y}$ для каждого класса. Оказывается, что необходимую информацию можно получить из изображения бернуллиевской сцены. Действительно, так как площадь проекции объекта известна, а размеры зоны интереса выбираются на этапе дешифрирования, то естественно определить $P(A)$ и $P(Q \setminus A)$ равенствами

$$P(A) = \frac{|A|}{|Q|}, \quad P(Q \setminus A) = \frac{|Q \setminus A|}{|Q|}.$$

Пусть x_Q – изображение зоны интереса, а $n_Q(y)$, $n_A(y)$ и $n_{Q \setminus A}(y)$ – количество пикселов с яркостью y , $y \in Y$, зоны интереса, объекта и фона соответственно, которые наблюдаются на x_Q , x_A и $x_{Q \setminus A}$. Известно, что относительные частоты

$$\begin{aligned} \bar{p}_Q(y) &= \frac{n_Q(y)}{|Q|}, \quad \bar{p}_A(y) = \frac{n_A(y)}{|A|} \\ \text{и } \bar{p}_{Q \setminus A}(y) &= \frac{n_{Q \setminus A}(y)}{|Q|}, \quad y \in Y, \end{aligned}$$

являются состоятельными оценками неизвестных вероятностей, образующих распределения

P_Q зоны интереса, P_A объекта и $P_{Q \setminus A}$ фона. Легко проверить, что они удовлетворяют уравнению

$$\bar{p}_Q(y) = P(A)\bar{p}_A(y) + P(Q \setminus A)\bar{p}_{Q \setminus A}(y), \quad y \in Y \quad (2)$$

с двумя неизвестными $\bar{p}_A(y)$ и $\bar{p}_{Q \setminus A}(y)$. Пусть $Fr(Q)$ – граница зоны интереса, $x_{Fr(Q)}$ – ее изображение и $n_{Fr(Q)}(y)$ – количество пикселов границы с яркостью y , $y \in Y$, на $x_{Fr(Q)}$. Очевидно, что $\bar{p}_{Fr(y)}(y)$, определяемая равенством

$$\bar{p}_{Fr(Q)}(y) = \frac{n_{Fr(Q)}(y)}{Fr(Q)}, \quad y \in Y,$$

так же, как и $\bar{p}_{Q \setminus A}(y)$, является оценкой для неизвестной $p_{Q \setminus A}(y)$. Из сходимости по вероятности оценок $\bar{p}_{Q \setminus A}(y)$ и $\bar{p}_{Fr(Q)}(y)$ к неизвестной вероятности $p_{Q \setminus A}(y)$ следует, что $\bar{p}_{Q \setminus A}(y) \approx \bar{p}_{Fr(Q)}(y)$ при большом объеме $|Fr(Q)|$ выборки. После замены в уравнении (2) неизвестной оценки $\bar{p}_{Q \setminus A}(y)$ оценкой $\bar{p}_{Fr(Q)}(y)$ оно превращается в уравнение с одним неизвестным $\bar{p}_A(y)$. Его решение завершает нахождение приближенных значений для вероятностей, образующих распределения объекта P_A и фона $P_{Q \setminus A}$, необходимые для классификации пикселов зоны интереса. При этом решающее правило h и вероятность ошибки $e(h)$ классификации принимают вид

$$\begin{aligned} h^{-1}(\{1\}) &= \{y \in Y : P(A)\bar{p}_A(y) \geq P(Q \setminus A)\bar{p}_{Q \setminus A}(y)\}, \\ h^{-1}(\{2\}) &= Y \setminus h^{-1}(\{1\}), \\ e(h) &= 1 - P(A) \times \\ &\times \sum_{y \in h^{-1}(\{1\})} \bar{p}_A(y) - P(Q \setminus A) \sum_{y \in h^{-1}(\{2\})} \bar{p}_{Q \setminus A}(y). \end{aligned}$$

Очевидно, что $e(h)$ с увеличением длины границы стремится к вероятности ошибки $e(h^*)$ байесовского (оптимального) решающего правила.

В изложенном методе сегментации по изображению $x_{Fr(Q)}$ границы зоны интереса вычислялись $|Y|$ оценок неизвестных вероятностей, образующих распределение $P_{Q \setminus A} = (p_{Q \setminus A}(y))_{y \in Y}$. При отсутствии выборки соответствующего объема, что является обычным делом в прикладных задачах, метод не применим. Однако для бернульиевских сцен можно воспользоваться центральной предельной теоремой. Действительно, пусть $\xi_A = (\xi_a)_{a \in A}$ – заданный объект с неизвестными средним значением m_A и дисперсией σ_A^2 , $B(\mathbf{z}, r)$ – круг, принадлежащий A , и \bar{x}_z – среднее арифметическое значение, вычисленное по кругу. При достаточно большом радиусе r распределение случайной величины \bar{x}_z можно считать нормальным с параметрами $(m_A, \sigma_A^2 / |B(\mathbf{z}, r)|)$, независимо от распределения x_z . Это позволяет вместо неизвестных распределений объекта P_A

и фона $P_{Q \setminus A}$ воспользоваться нормальными распределениями с параметрами $(\bar{x}_A, \sigma_A^2 / |B(\mathbf{0}, r)|)$ и $(\bar{x}_{Q \setminus A}, \sigma_{Q \setminus A}^2 / |B(\mathbf{0}, r)|)$ соответственно. В качестве неизвестных параметров $\bar{x}_{Q \setminus A}$ и $\sigma_{Q \setminus A}^2$ фона используются их оценки \bar{x}_{Fr} и σ_{Fr}^2 , вычисленные по изображению границы:

$$\begin{aligned} \bar{x}_{Fr} &= \frac{1}{|Fr(Q)|} \sum_{\mathbf{z} \in Fr(Q)} x_{\mathbf{z}}, \\ \sigma_{Fr}^2 &= \frac{1}{|Fr(Q)|} \sum_{\mathbf{z} \in Fr(Q)} (x_{\mathbf{z}} - \bar{x}_{Fr})^2. \end{aligned}$$

Неизвестные параметры объекта \bar{x}_A и σ_A^2 определяются из уравнений

$$\bar{x}_Q = P(A)\bar{x}_A + P(Q \setminus A)\bar{x}_{Q \setminus A} \quad (3)$$

и

$$\sigma_Q^2 + \bar{x}_Q^2 = P(A)(\sigma_A^2 + \bar{x}_A^2) + P(Q \setminus A)(\sigma_{Q \setminus A}^2 + \bar{x}_{Q \setminus A}^2) \quad (4)$$

после замены в них неизвестных $\bar{x}_{Q \setminus A}$ и $\sigma_{Q \setminus A}^2$ приближенными значениями, вычисленными по границе зоны интереса. Напомним, что в качестве признака пикселя $\mathbf{z} \in Q \setminus Fr(Q)$ требуется рассматривать не его изображение $x_{\mathbf{z}}$, а оценку $\bar{x}_{\mathbf{z}}$ неизвестного среднего значения $E\xi_{\mathbf{z}}$.

Обобщение байесовских методов на случай нескольких изображений ($v > 1$) очевидно. Так как количество оцениваемых вероятностей экспоненциально зависит от v , то их непосредственная оценка в прикладных задачах представляется весьма проблематичной. При использовании многомерного варианта центральной предельной теоремы по выборке достаточно оценить v координат вектора средних значений и $v(v-1)$ ковариаций.

Компьютерные эксперименты

Для иллюстрации изложенных методов сегментации зон интереса проведены компьютерные эксперименты. В них используются фрагменты трех пространственно совмещенных изображений реальной сцены, полученные с летательного аппарата в спектральных зонах 0,7–1,1, 3,0–5,0 и 8,0–12,0 мкм. Каждый пиксель имеет форму квадрата со стороной 0,3 м. На каждом фрагменте программным способом построены изображения восьми воображаемых объектов, отсутствующих на реальной сцене. Предполагается, что воображаемые объекты имеют прямоугольную форму, случайные координаты и ориентацию, а их размеры совпадают с размерами реальных объектов. Яркости пикселов каждого воображаемого объекта являются независимыми в совокупности случайными величинами с нормальным законом распределения. Его парамет-

Свойства сцены (отношение сигнал/шум)

Номер изображения	Номера объектов							
	1	2	3	4	5	6	7	8
1	2,79	1,60	4,68	2,34	2,0	5,51	3,40	2,60
2	3,29	5,79	0,42	1,23	1,70	1,00	0,22	0,77
3	7,13	6,66	0,58	1,72	2,84	1,41	1,87	2,20

рами служат оценки среднего значения и дисперсии реальных объектов. Для первого фрагмента $m_A^1 = 46$ и $(\sigma_A^1)^2 = 214$, для второго $m_A^2 = 220$ и $(\sigma_A^2)^2 = 104$, для третьего $m_A^3 = 230$ и $(\sigma_A^3)^2 = 27$. Свойства сцены в окрестности каждого объекта описываются отношением сигнал/шум k

$$k = \frac{|m_A - m_{Q|A}|}{\sigma_A + \sigma_{Q|A}}$$

Его значения для каждого изображения приведены в таблице. В качестве примера на рис. 1 приведен фрагмент первого (лучшего, с точки зрения отношения сигнал/шум) изображения сцены с минимальным значением $k = 1,60$.

Каждое изображение сцены использовалось вначале для поиска зон интереса, а затем – для их сегментации. Результаты сегментации зон интереса методом квантилей с применением только первого изображения представлены на рис. 2, а с применением всех трех изображений – на рис. 3. Сравнение полученных результатов позволяет отметить два важных обстоятельства. Во-первых, построить зону интереса для второго объекта с использованием только первого изображения не удалось. Во-вторых, координаты зон интереса, выявленных с использованием только первого и с использованием всех трех изображений, оказались разными.

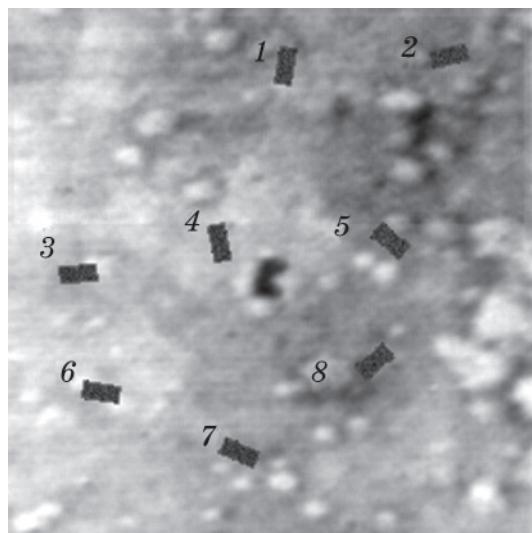


Рис. 1. Первое изображение сцены.

Если сцена описывается бернуллиевским случайнм полем, то изображение каждого объекта является случайной выборкой и может использоваться для оценки вероятностей и числовых параметров случайных яркостей. На рис. 4 представлены результаты байесовской классификации пикселов зон интереса с оценкой неизвестных вероятностей, образующих

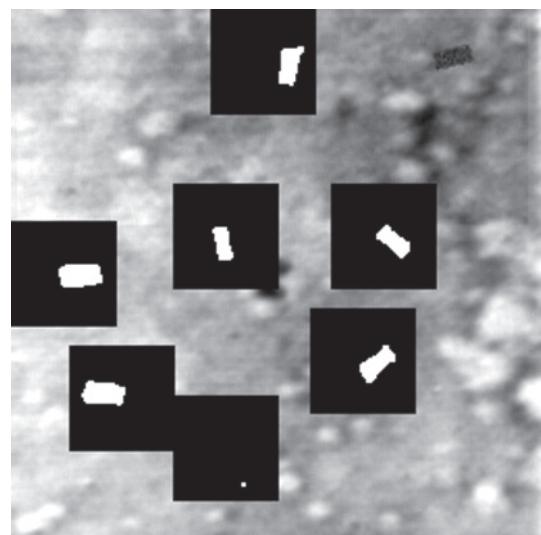


Рис. 2. Результаты сегментации по первому изображению.

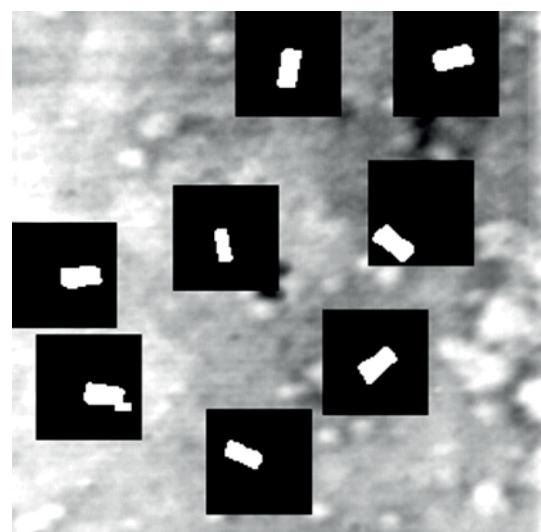


Рис. 3. Результаты сегментации по трем изображениям.

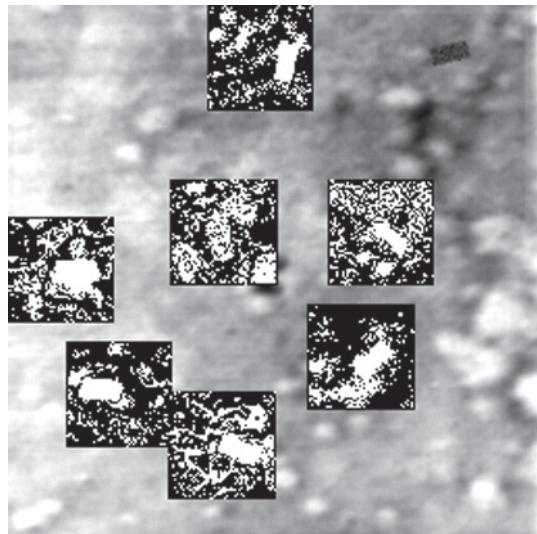


Рис. 4. Байесовская классификация. Общий случай.

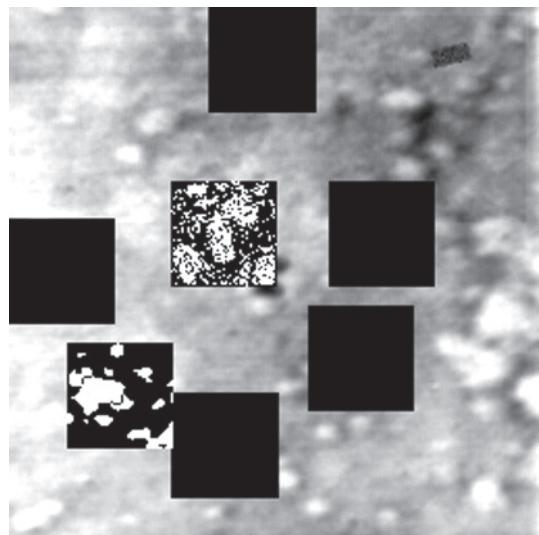


Рис. 5. Байесовская классификация. Оценка параметров.

распределения яркостей объекта и фона (общий случай).

На рис. 5 представлены результаты байесовской классификации в случае, когда в качестве признака пикселя $z \in Q \setminus Fr(Q)$ использовалось среднее арифметическое значение \bar{x}_z , получаемое осреднением по кругу. В соответствии с центральной предельной теоремой распределение такой суммы стремится к нормальному при увеличении радиуса. Черные квадраты на изображении означают, что в соответствующих зонах интереса оценки для неизвестных параметров \bar{x}_A и s_A^2 , полученные в результате решения уравнений (3) и (4), не удовлетворяют очевидным

ограничениям. Действительно, при сегментации в этих зонах оценки для \bar{x}_A оказались слишком большими (больше чем $|Y| - 1 = 255$), а оценки для дисперсии s_A^2 – даже отрицательными. Причина появления таких значений заключается в том, что изображения зон интереса рассматриваемой сцены не являются случайными выборками, т. е. исходные предположения о свойствах сцены оказались ложными.

Заключение

В работе рассматриваются методы сегментации зон интереса для двух моделей реальной сцены. Для локально изотропных сцен с корреляционной функцией, стремящейся к нулю, единственным признаком объекта, вычисляемым по изображению, является оценка его среднего значения. Для бернульиевских сцен по изображению возможно (по крайней мере теоретически) восстановление распределения вероятностей объекта. Это позволяет построить для сегментации зоны интереса асимптотически оптимальное решающее правило.

Тем не менее результаты компьютерных экспериментов по сегментации зон интереса рассмотренной в настоящей работе сцены показывают, что применение байесовских (оптимальных) методов может дать ложные результаты. Подобная ситуация возникает в тех случаях, когда используемые при построении математической модели предположения для реальной сцены не выполняются.

ЛИТЕРАТУРА

1. Фофанов В.Б. О теоретико-вероятностной формализации задачи дешифрирования аэрокосмических изображений // Автометрия. 2003. № 6. С. 107–118.
2. Яглом А.М. Корреляционная теория стационарных случайных функций. Л.: Гидрометеоиздат, 1981. 280 с.
3. Фофанов В.Б. Случайное расстояние и его применения в дешифрировании изображений // Материалы IX Междунар. конф. "Интеллектуальные системы и компьютерные науки". М.: изд-во ММФ МГУ, 2006. Т. 2. Ч. 2. С. 286–288.
4. Фофанов В.Б., Демченко А.В., Кулев Р.Ф. Дешифрирование многозональных изображений: методы и результаты // Оптический журнал. 2007. Т. 74. № 3. С. 55–59.
5. Тутубалин В.Н. Теория вероятностей. Краткий курс и научно-методические замечания. М.: изд-во МГУ, 1973. 232 с.